

•	Ser TCC	Met ATG	₽	Asp GAT	శ్ర్ క్ర	Val GTA
	¥ SA ACA	¥CG	8 8	₹ 8	Ser TCT	Lys AAG
-107 Met \ GTG	A Se	AGC Ser	Val GTA	∌ §	₹ S	₹
WAG.	æ 89	60 Met ATG	TAT	ot. Val GTT	Lee CTG	Asp GAT
3 <u>6</u> 47	를 크	₽ Ş	₹₹	T. Y.	Ala GCT	6 g 2 C T C C T C C T C C T C C C C C C C C
RBS	SC Aa	ల్ న	15 15	Ata GCT	25 25	His CAT
AAAĞ	Met	8₹	క్రక్తి	val GTC	SC AB	Ser
AAAA	ACG.	#E	Lys AG	Ser AGC	Lys AA	Ser TCT
WATG	#E	Gly GGG	8 g 8	Pro CCG	le ATT	Asp GAT
RBS iattctgcaaatgaaaaag <u>gagagg</u> ataaaga	90 le ATC	val GTC	val GTG	Asp	⊊ క్ క్రే	lle ATC
TATT	TA Le	IS ATT	₹\$	eys AA	Ş€ 1C Y	G √ GGT
4	Aa	PRO Tyr Be TAT AT	G6G G6G	Lys AA	Val GTA	Ser AGC
) TCTA1	Lee TA	Lys AA	ე მჭ	teu TTG	ე ლ	Asp GAC
P ATATTATTCCA <u>IACIAT</u> ACAATTAATACACAGAATAATCTGTCTAT	Ala	-70 Lys AAG	₹\$.20 Glu GAA	Tyr TAC	lle ATC
	# E	∂ A	⊒ Se	Lys AAA	SC 38	30 Val GTT
	۳ ع و و ق	09 x	Ser TCT	Val GTA	Val GTG	≱ ₽
	PRE Leu Le	Asn	lle ATT	Ala GCT MAT	S €	Val GTA
⊕+¥ ¥±	Ser	₹ 1Ç	Val GTC	₹ ₹	GP CAG	Lys AA
AŢĀC	ATC	र ≸	& & &	3€	Aga - GCG	Val GTT
AŢĀĊĪ	₽ 100	G G	₹\$	-54 AC	- ₹- 14C	Asn
P ST <u>a</u> ctaaaatattattcc/	Val	₽ 20	Lys AAG	TA E	≱ 8 GCG	శ్రే శ్ర
	Lys AAA	Ala GCG	Lys AAG	ACA	His CAT	∂ 86
	r A	-80 Ala Gin GCC CAG	Ala GCT	Ala GCT	SC AB	ACT ACT
	0 6	S & &	As GCC	ક્ષ્ટ્રે કુ	Val GTA	Ty TAC
© <u>→</u>	AGA	ici Se	AGC A	Ala GCT	₹ S	ද දීල්
_	86	174	249	324	333	474

FIG._1B-1

Ala GCC	Lys AA	Mei	Ala GCA	Gy GGT	Pro CCT	GG.	ACT
val GTT	Val GTA	Asn	Val GTT	55 7			Asn .
CA C	Ala GCT	Asn	Ala GCC	TY TAC	Val GTA	Ty TAC	ACA
ACT ACT	TAC TAC	₽ ₹	8.5€	£ 88€	Se AGC		TG 2
Gy GGA	S is E	₽ ATC	140 GAT	val GTG	<u> </u>		240 Asn AAC
SAC CAC	Z ≤ 45 Z ≤ 45	Ala GCG	Val GTT	₽ Ş	± 3€	TAC TAC	£ 933
Ser TCT	Ser SCA SCA	Тр ТGG	Ala GCA	Ser AGC	Ser TCT	₹₹	His
Asn	Ser AGC	Glu GAG	Ala GCG	Ser TCA	₽ QC ¥9	Asn	Lys
Asp Asn AAC	₽ 23	lle ATC	ξ. \$	Ser AGC	Arg AGA		Ser TCT
60 Asp GAC	Ata GCG	ე გე	E TA	35 SG SG SG	S G		CH CH
S S	Val GTT	Asn	Ala GCT	Se ⊒ TCC	Asn	CT CT	e H
Phe TTC	299 2 5 0	lle A∏	Ala GCT	Ser Thr ACT	Ser AGC		Le I
Asn Pro CCT	∃ E	lle ATC	Ser TCT	ე ლ <u>ჯ</u>	Ser AGC	Ser . AGC	Ala GCT D
Pro Asn AAT	val GTA	1ր 16 G	Gly GGT	Glu GAA	Asp GAC		Ala GCT
Thr	8 65,	Ser AGC	130 Ser TCT	Asn	180 Val GTT	lle ATC	230 Ala Ala A GCG GCT FIG 1
o Serie	lle ATC	Ty TAC	Pro CCT	Gly GGT	Ala GCT		₹ 86.64
Ser TCT	Ser TCA	S G	GGA GGA	SCC CCC	ე ცე	Val GTA	Ata GCC
P. CC1	Asn	ခွဲ့ ဗိ	ე დ_ _	₽ 25	Val GTA	ر وو خ	Val GTT
Val GTT	Asn	15 6	CTC	Ata GCG	₽ ₽	Po CCT	SA CAC
50 Met	E CH	<u>6</u> 64 54	Ser AGC	150 Val GTT	Re ATT	8 8 8 8 8 8 8	7 200
Ser AGC	Ala GCT	Ata Asp GAC	Met ATG	Val GTC	val GTC	Mei ATG	Ser TCT
Ala GCC	A B 600	Asp Ata GCT	Asn	Val GTA	Se TCT	Val GTC	Ala GCA
GGA GGA	val GTT	6 <u>4</u> 661	A T	val GTC	Pro CCT	Asp GAT	Met ATG
09¢ 09¢	ACA T	SE CIC	Val GTT	ر و و	TAC TAC	æ Lo	કુ <u>ર</u> ૂ
Ala GCA	26 66 66 66	Val GTT	128 GAC GAC	Se 100	5 7 §	Se GAG	220 Thr
549	624	669	774	849	924	666	1074

250 Gin Gin Val Arg Ser Ser Leu Glu Asn Thr Thr Lys Leu Gly Asp Ser Phe Tyr Tyr Gly Lys Gly Leu lie Asn 1149 CAA GTC CGC AGC AGT TTA GAA AAC ACC ACT ACA AAA CTT GGT GAT TCT TTC TAC TAT GGA AAA GGG CTG ATC AAC

270

Val Gin Ala Ala Gin OC

1224 GTA CAG GCG GCA GCT CAG TAA AACATAAAAAACGGGCCTTGGCCCCGGCGTTTTTTATTTTTCTTCCTCCGCATGTTCAATCCGCTCC

1316 ATAATGACGGATGGCTCCCTCTGAAAATTTTAACGAGAAACGGCGGGTTGACCGGCTCAGTCCCGTAACGGCCAAGTCCTGAAACGTCTCAATCGCCG

1416 CTTCCCGGTTTCCGGTCAGCTCAATGCCGTAACGGTCGGCGGCGTTTTCCTGATACCGGGGAGACGGCATTCGTAATCGGATC

FIG._1B-3

FIG._1B-3 FIG._1B-2 FIG._1B - 1 FIG._1B

CONSERVED RESIDUES IN SUBTILISINS FROM BACILLUS AMYLOLIQUEFACIENS AQSVP.G...APA.H..G 21 . TGS. VKVAV. D.G....HP DL . . . GGAS . VP QD . N . HGTHVAGT . AALNNSIG 90 100 V L G V A P S A . L Y A V K V L G A . G 130 V.N.SLG.PS.S....A.. · · · · · G V · V V A A · G N · G · · · 180 · · · · · · Y P · · Y · · · · A V G A · 200 D. N. ASFS. G. LD. A PGV..QST.PG..Y...NGT 230 SMA.PHVAGAAAL...K... 241 250 W . . . Q . R . . L . N T . . . L G Y G . G L . N . . A A . .

FIG._2

COMPARISON OF SUBTILISIN SEQUENCES FROM:
B.amyloliquefaciens
B.subtilis
B.licheniformis
B.lentus

4444 HH លលល* SSAF 9 9 9 W 0000 SSEE 999 2 2 4 2 4 4 2 4 4 4 K A A > > > > > z z Z 02 K 03 000 **FFK** × × F 니 000 000 000 9992 0 0 K Z HHO 4464 **4 × 4** 4404 4 44 001K HHHV KKK0 ខ្លួ **>** H H H Ö ×× 4 > > > တလ E S 0000

O O O Oннын E S ZOZ コココ KK **444 P** H + H O O 0 0 4 ** > > > Ξ H H H H H 0000 **HHHH** 8 8 9 9 ZSZZ Z 000 Ω 999 аана * 4 14 Д A 2 0 ZXA H H K * ы 田田 **හ** හ හ හ 4 A 50 M V > > Δ, 8 8 8 8 AAA 0000 00 00 4 **x** > x > >

999 XXXX 200 zz KWHZ 4 ⋖ KK 333 医医医医 н нл Ç 000 zzsa H H > K нннн BBBBB 0 0 0 0 4444 9 9 8 8 Ç 000 Ø တတတ O 000 2 5 8 8 **4 55 55 4** 0 2 2 0 コレコロ > >> KK K > > 90 LYA 4 ~ LY > > a a ល ល ល ស < **A > A** Ø 20 00 00 Д 4 4 Ø O 000 H 81 V 1

0000 **ಬ** ಬ ಬ ಬ 0000 ではららら 2222 0000 ** < ~ K K VVA 4 Ø VVA **>** > > > H > 0000 S S R R K S K 03 CAAVDKAV CTVVDKAV CQAVDNAY CQAVNSATE KKKE JX 4 4 F A A 0 0 0 0 0000 0 H 0 0 444 0000 000 444 0 0 0 0 EEEG z zz ΙΛ н н «

FIG._3A

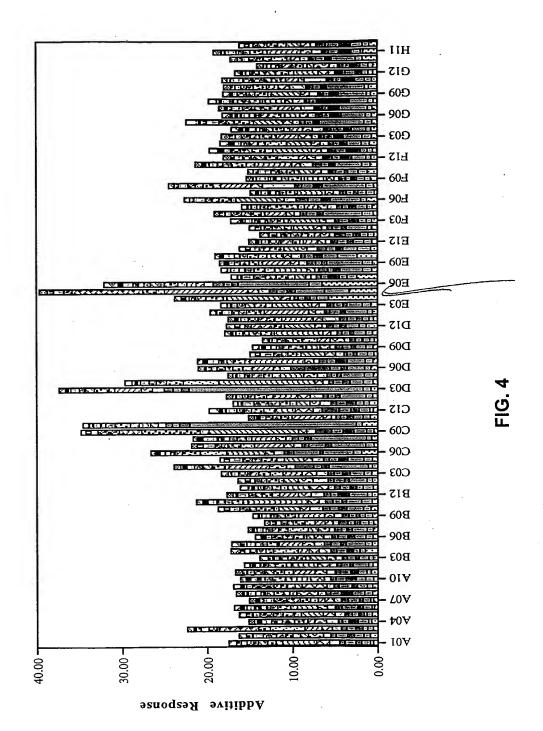
444 x > 2222 227 医医鼠氏 S & & 0000 > < > > 8 8 8 8 ល ល ល 444 8 8 8 8 ** **8 8 8 8** aazz ZZSZ SSZZ 8888 > > > F 0 X Q Q ** ပ ပ ပ ပ >>>> ខេខខ 404 DAAA 4 4 *** 9 0 H H H * 8 8 8 8 8 161 S S S S T S +

ZFZG **HHHZ** KKKK 8 8 8 9 とし ひた ннн> 2222 ** 230 A A 4 4 4 4 4 4 0000 ** >>> **###** 444 8 H 8 H × 4 A A ZZZZ 220 11 S 11 S 11 S 0000 2222 ы **44** F 0 **** **KEEE** 2020 210 P G P T よよで で ** 8 8 8 8 0400 H H > > 8 8 8 8 201 P G V P G A

9992 KKKH ** 270 V Q A V C A A E A нннь ココココ 0000 K K K S 000 **** × アエロ 260 S S F T F F OZSS 000 227 *** FAKA *** z o o z T I I I I I 3段段田 2222 **K K K K** > > > H aaaaH K 03 > ZZKZ

FIG._3B

FIG..3A FIG..3B



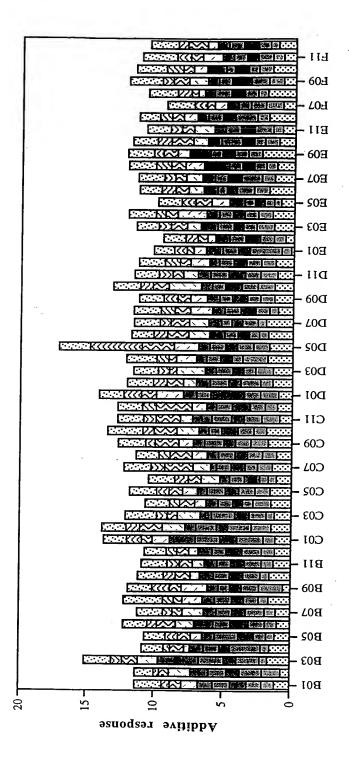


FIG. €

		•	•		•
1	A12	IKDFHVYFRESRDAG	49	E12	SATSRGVLVVAASGN
2		LEQAVNSATSRGVLV			
	A11		50	E11"	SRGVLVVAASGNSGA
3	A10	AQSVPWGISRVQAPA	51	E10	VLVVAASGNSGAGSI
4	Α9	VPWGISRVQAPAAHN	52	E9	VAASGNSGAGSISYP
5		GISRVQAPAAHNRGL			
	A8		53	E8	SGNSGAGSISYPARY
6	Α7	RVQAPAAHNRGLTGS	54	E7	SGAGSISYPARYANA
7	Α6	APAAHNRGLTGSGVK	55	E6	GSISYPARYANAMAV
8		AHNRGLTGSGVKVAV	56		
	A5			E5	SYPAR <u>YANAMAVGA</u> T
9	Α4	RGLTGSGVKVAVLDT	57	E4	<u>ARYANAMAVGA</u> TDON
10	A3	TGSGVKVAVLDTGIS	58	E3	ANAMAVGATDONNNR
11		GVKVAVLDTGISTHP	59		
12	A2	-		E2	MAVGATDQNNNRASF
	A1	VAVLDTGISTHPDLN	60	E1	GATDQNNNRASFSQY
13	B12	LDTGISTHPDLNIRG	61	F12	DQNNNRASFSQYGAG
14		GISTHPDLNIRGGAS	62		
15	B11			F11	NNRASFSQYGAGLDI
	B10	THPDLNIRGGASFVP	63	F10	ASFSQYGAGLDIVAP
16	B9	DLNIRGGASFVPGEP	64	F9	SQYGAGLDIVAPGVN
17	B8	IRGGASFVPGEPSTQ	65		
18				F8	GAGLDIVAPGVNVQS
	В7	GASFVPGEPSTQDGN	66	F7	LDIVAPGVNVQSTYP
19	B6	FVPGEPSTQDGNGHG	67	F6	VAPGVNVQSTYPGST
20	B5	GEPSTQDGNGHGTHV	68	F5	GVNVQSTYPGSTYAS
21					
	B4	STQDGNGHGTHVAGT	69	F4	VQSTYPGSTYASLNG
22	В3	DGNGHGTHVAGTIAA	70	F3	TYPGSTYASLNGTSM
23	B2	GHGTHVAGTIAALNN	71	F2	GSTYASLNGTSMATP
24		THVAGTIAALNNSIG	72		
25	B1			F1	YASLNGTSMATPHVA
	C12	AGTIAALNNSIGVLG	7 3	G12	LNGTSMATPHVAGAA
26	C11	IAALNNSIGVLGVAP	74	G11	TSMATPHVAGAAALV
27	C10	LNNSIGVLGVAPSAE	75		ATPHVAGAAALVKOK
28		SIGVLGVAPSAELYA		G10	
	C9		76	G9.	HVAGAAALVKQKNPS
29	С8	VLGVAPSAELYAVKV	77	G8-	GAAALVKQKNPSWSN
30	C7	VAPSAELYAVKVLGA	78	G7	ALVKQKNPSWSNVQI
31.	C6	SAELYAVKVLGASGS	79		
32				Gб	KOKNPSWSNVQIRNH
	C5	LYAVKVLGASGSGSV	80	G5	NPSWSNVQIRNHLKN
33	C4	VKVLGASGSGSVSSI	81	G4	WSNVQIRNHLKNTAT
34	C3	LGASGSGSVSSIAQG	82		VQIRNHLKNTATSLG
35:				G3	
	C2	SGSGSVSSIAQGLEW	83	G2	RNHLKNTATSLGSTN
36:	C1	GSVSSIAQGLEWAGN	84	G1	LKNTATSLGSTNLYG
37	D12	SSIAQGLEWAGNNGM	85	H12	TATSLGSTNLYGSGL
38		AQGLEWAGNNGMHVA			
	D11		86	H11	SLGSTNLYGSGLVNA
39	D10	LEWAGNNGMHVANLS	87	H10	STNLYGSGLVNAEAA
40	D9	AGNNGMHVANLSLGS	88	Н9	NLYGSGLVNAEAATR
41	D8	NGMHVANLSLGSPSP		1	THE TOO CE VIN IN THE TEN
42					
	D7	HVANLSLGSPSPSAT			
43	D6	NLSLGSPSPSATLEQ			
44	D5	LGSPSPSATLEQAVN			
45					
	D4	PSPSATLEQAVNSAT			
46	D3	SATLEQAVNSATSRG			
47	D2	LEQAVNSATSRGVLV			•
48	D1	AVNSATSRGVLVVAA			
-	ΝŢ	TIDITIDIO VIDA AUV			•

FIG. 6A

		T.//D.C./// T.// C.				
1	A12	IKDFHVYFRESRDAG		49	E12	KKIDVLNLSIGGPDF
2	A11	DAELHIFRVFTNNQV		50	E11	DVLNLSIGGPDFMDH
3	A10	PLRRASLSLGSGFWH		51	E10	NLSIGGPDFMDHPFV.
4	A9	RASLSLGSGFWHATG		52	E9	IGGPDFMDHPFVDKV
5	A8	LSLGSGFWHATGRHS		53	E8 -	PDFMDHPFVDKVWEL
6	A7	GSGFWHATGRHSSRR		54	E7	MDHPFVDKVWELTAN
7	A6	FWHATGRHSSRRLLR		55	E6	PFVDKVWELTANNVI
8	A5	ATGRHSSRRLLRAIP	,	56	E5	DKVWELTANNVIMVS
9	A4	RHSSRRLLRAIPRQV		57	E4	WELTANNVIMVSAIG
10	А3	SRRLLRAIPRQVAQT		58	E3	TANNVIMVSAIGNDG
11	A2	LLRAIPRQVAQTLQA		59	E2	NVIMVSAIGNDGPLY
12	A 1	AIPRQVAQTLQADVL		60	E1	MVSAIGNDGPLYGTJ.
13	B12	RQVAQTLQADVLWOM		61	F12	AIGNDGPLYGTLNNP
14	B11	AQTLQADVLWOMGYT		62	F11	NDGPLYGTLNNPADO
15	B10	LQADVLWQMGYTGAN		63	F10	PLYGTLNNPADOMDV
16	B9	DVLWQMGYTGANVRV		64	F9	GTLNNPADQMDVIGV
17	B8	WQMGYTGANVRVAVF		65	F8	NNPADOMOVIGVGGI
18	B7	GYTGANVRVAVFDTG		66	F7	ADOMDVIGVGGIDFE
19	B6	GANVRVAVFDTGLSE		67	F6	MDVIGVGGIDFEDNI
20	B5	VRVAVFDTGLSEKHP		68	F5	IGVGGIDFEDNIARF
21	B4	AVFDTGLSEKHPHFK		69	F4	GGIDFEDNIARFSSR
22	В3	DTGLSEKHPHFKNVK		70	F3	DFEDNIARFSSRGMT
23	B2	LSEKHPHFKNVKERT		71	F2	
24	B1	KHPHFKNVKERTNWT		72	F1	DNIARFSSRGMTTWE
25	C12	HFKNVKERTNWTNER		73	G12	ARFSSRGMTTWELPG
26	C11	NVKERTNWTNERTLD			G11	SSRGMTTWELPGGYG
27	C10	ERTNWTNERTLDDGL		74	G10	GMTTWELPGGYGRMK
28	C9	NWTNERTLDDGLGHG		75 76		TWELPGGYGRMKPDI
29	C8	NERTLDDGLGHGTFV		76	G9	LPGGYGRMKPDIVTY
30	. C7			77	G8	GYGRMKPDIVTYGAG
31	C6	TLDDGLGHGTFVAGV		78	G7	RMKPDIVTYGAGVRG
32	C5	DGLGHGTFVAGVIAS		79	G6	PDIVTYGAGVRGSGV
33	C4	GHGTFVAGVIASMRE		80	G5	VTYGAGVRGSGVKGG
34	C3	TFVAGVIASMRECQG		81	G4	GAGVRGSGVKGGCRA
35	C2	AGVIASMRECQGFAP		82	G3	VRGSGVKGGCRALSG
36		IASMRECQGFAPDAE		83	G2	SGVKGGCRALSGTSV
37	C1 D12	MRECQGFAPDAELHI		84	G1	KGGCRALSGTSVASP
		CQGFAPDAELHIFRV		85	H12	CRALSGTSVASPVVA
38 39	D11	FAPDAELHIFRVFTN		86	H11	LSGTSVASPVVAGAV
	D10	DAELHIFRVFTNNQV		87	H10	TSVASPVVAGAVTLL
40	D9	LHIFRVFTNNQVSYT		88	Н9	ASPVVAGAVTLLVST
41 42	D8	FRVFTNNQVSYTSWF		89	Н8	VVAGAVTLLVSTVQK
	D7	FTNNQVSYTSWFLDA	,	90	H7	GAVTLLVSTVQKREL
43	D6	NQVSYTSWFLDAFNY		91	Н6	TLLVSTVQKRELVNP
44	D5	SYTSWFLDAFNYAIL	•	92	H5	VSTVQKRELVNPASM
45	D4	SWFLDAFNYAILKKI		93	H4	VQKRELVNPASMKQA
46	D3	LDAFNYAILKKIDVL		94	Н3	RELVNPASMKQALIA
47	D2	FNYAILKKIDVLNLS		95	H2	VNPASMKQALIASAR
48	D1	AILKKIDVLNLSIGG	•	96	H1	ASMKQALIASARRLP

FIG. 6B

97	I12	IKDFHVYFRESRDAG
98	I11	DAELHIFRVFTNNOV
99	I 10 '	KQALIASARRLPGVN
100	19	LIASARRLPGVNMFE
101	18	SARRLPGVNMFEOGH
102	I7	RLPGVNMFEQGHGKL
103	16	GVNMFEQGHGKLDLL
104	15	MFEQGHGKLDLLRAY
105	· 14	QGHGKLDLLRAYOIL
106	. I3	GKLDLLRAYOILNSY
107	12	DLLRAYQILNSYKPQ
108	I1	RAYQILNSYKPQASL
109	J12	QILNSYKPQASLSPS
110	J11	NSYKPQASLSPSYID
111	J10	KPQASLSPSYIDLTE
112	J9	ASLSPSYIDLTECPY
113	J8	SPSYIDLTECPYMWP
114	J7	YIDLTECPYMWPYCS
115	J6	LTECPYMWPYCSQPI
116	J5	CPYMWPYCSQPIYYG

FIG. 6C

MKLVNIWLLLLVVLLCGKKHLGDRLEKKSFEKAPCPGCSHLTLKVEFSSTVVEYEYIVAFNGYFT AKARNSFISSALKSSEVDNWRIIPRNNPSSDYPSDFEVIQIKEKQKAGLLTLEDHPNIKRVTPQR KVFRSLKYAESDPTVPCNETRWSQKWQSSRPLRRASLSLGSGFWHATGRHSSRRLLRAIPRQVAQ TLQADVLWQMGYTGANVRVAVFDTGLSEKHPHFKNVKERTNWTNERTLDDGLGHGTFVAGVIASM RECQGFAPDAELHIFRVFTNNQVSYTSWFLDAFNYAILKKIDVLNLSIGGPDFMDHPFVDKVWEL TANNVIMVSAIGNDGPLYGTLNNPADQMDVIGVGGIDFEDNIARFSSRGMTTWELPGGYGRMKPD IVTYGAGVRGSGVKGGCRALSGTSVASPVVAGAVTLLVSTVQKRELVNPASMKQALIASARRLPG VNMFEQGHGKLDLLRAYQILNSYKPQASLSPSYIDLTECPYMWPYCSQPIYYGGMPTVVNVTILN GMGVTGRIVDKPDWQPYLPQNGDNIEVAFSYSSVLWPWSGYLAISISVTKKAASWEGIAQGHVMI TVASPAETESKNGAEQTSTVKLPIKVKIIPTPPRSKRVLWDQYHNLRYPPGYFPRDNLRMKNDPL ${\tt DWNGDHIHTNFRDMYQHLRSMGYFVEVLGAPFTCFDASQYGTLLMVDSEEEYFPEEIAKLRRDVD}$ NGLSLVIFSDWYNTSVMRKVKFYDENTRQWWMPDTGGANIPALNELLSVWNMGFSDGLYEGEFTL ANHDMYYASGCSIAKFPEDGVVITQTFKDQGLEVLKQETAVVENVPILGLYQIPAEGGGRIVLYG DSNCLDDSHRQKDCFWLLDALLQYTSYGVTPPSLSHSGNRQRPPSGAGSVTPERMEGNHLHRYSK VLEAHLGDPKPRPLPACPRLSWAKPQPLNETAPSNLWKHQKLLSIDLDKVVLPNFRSNRPQVRPL SPGESGAWDIPGGIMPGRYNQEVGQTIPVFAFLGAMVVLAFFVVQINKAKSRPKRRKPRVKRPQL MQQVHPPKTPSV

FIG. 7

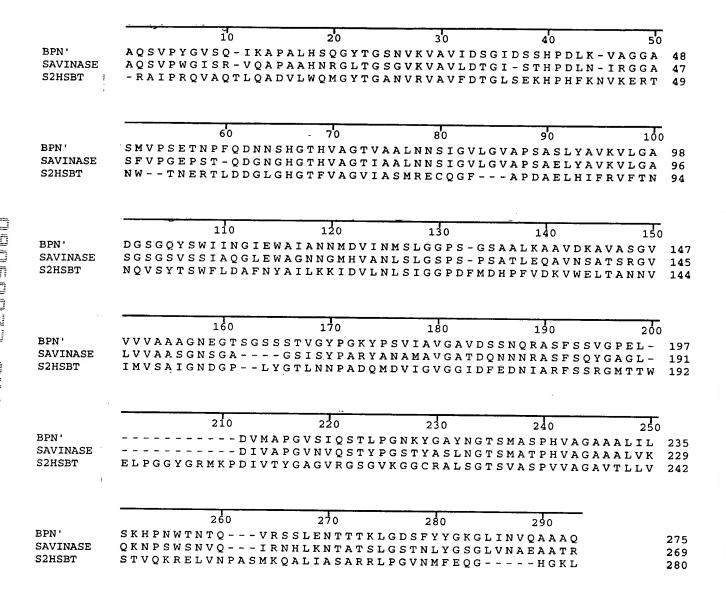
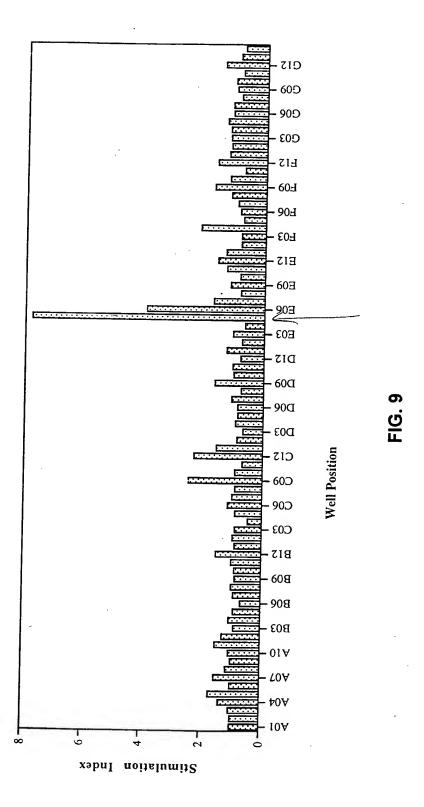


FIG. 8



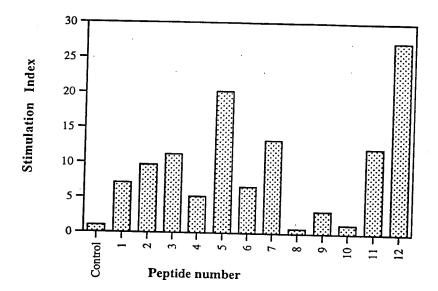


FIG. 10